Serial ATA and Serial Attached SCSI technologies

technology brief

Abstract 2

Introduction 2

Serial architecture: the future of HDD technologies 2
- Parallel ATA technology 2
- Parallel SCSI 4
- Serial drive technologies 5
  - LVD signaling 5
  - Point-to-point connections 6
  - 8b/10b encoding 6

Serial ATA technology 7
- SATA 1.0 7
  - Key components 8
- SATA II 9
  - Performance enhancements 9
  - Backward compatibility 9
- SATA positioning and roadmap 10

Serial Attached SCSI technology 10
- Performance 10
- Key components 11
- SAS/SATA compatibility 11
- SAS topologies 12
  - Internal 12
  - Internal/external with JBODs 12
  - Multi-node clusters 13
- SAS positioning and roadmap 13

Conclusion 14

Call to action 15

Abstract

This technology brief explains the technical differences between parallel ATA and SCSI and their serial counterparts—Serial ATA and Serial Attached SCSI. It also explains why the industry is migrating to these serial technologies to meet the bandwidth requirements of future enterprise technologies.

Introduction

IT managers have limited flexibility when choosing a data center storage solution. Their options have been limited to separate systems based on parallel Advanced Technology Attachment (ATA), Small Computer Systems Interface (SCSI), or Fibre Channel disk interfaces. These storage options limit the IT manager's ability to deploy and redeploy core technologies, thereby adding significant cost and management burden to the enterprise. With the advent of serial I/O architectures, key solution providers like HP can now provide one system or storage solution that will meet the requirements for a broad range of storage applications. This will allow IT managers to standardize on a single server or storage platform, which can dramatically reduce the complexity of managing the assortment of storage applications in today's data center.

HP has actively participated in the Serial ATA (SATA) and Serial Attached SCSI (SAS) industry working groups to develop specifications for serial I/O technologies that will provide the flexibility and performance IT managers need for storage solutions. Serial ATA will be implemented in desktop PCs and low-cost, non-mission-critical server storage solutions. Serial Attached SCSI will be implemented in enterprise class devices providing the full functionality of SCSI hard disk drives (HDDs).

This paper describes the current state of parallel ATA and SCSI technologies and the barriers they face in meeting the requirements of future enterprise technologies. Then it describes the fundamental features of serial I/O technologies that enable them to overcome the limitations of these parallel I/O technologies. Subsequent sections describe SATA and SAS technologies and the applications in which they will be deployed.

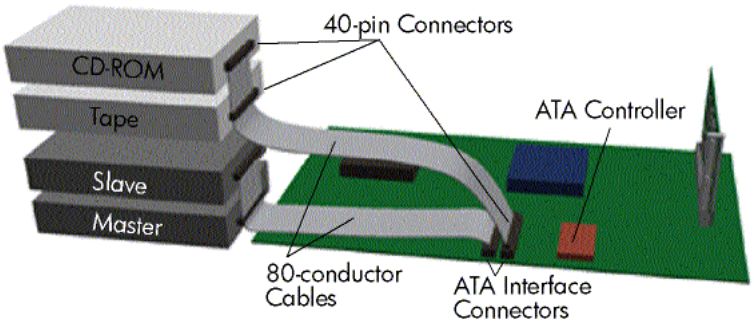
Serial architecture: the future of HDD technologies

Today's dominant parallel I/O architectures, ATA and SCSI, face several challenges in meeting the requirements of future enterprise technologies. This section describes the limitations of parallel ATA and SCSI to meet future enterprise requirements and why the industry is migrating toward serial I/O technologies.

Parallel ATA technology

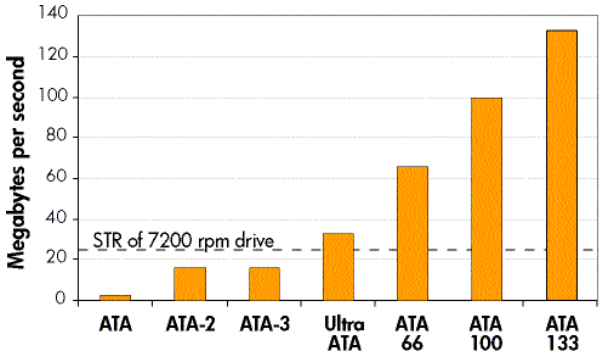
Parallel ATA has been the dominant interface for desktop and notebook products since it was introduced the 1980's as IDE (Integrated Drive Electronics). Figure 1 shows the major components in the parallel ATA architecture. An ATA controller and two parallel ATA connections are built into the motherboard. Up to two devices can be connected to each ATA connector by using an 80-conductor ribbon cable with three 40-pin connectors (26 signal pins and 15 command pins). The bulky cable impedes airflow in the cabinet, which is increasingly important for cooling; however, it is adequate for small personal systems with less than four drives.

Figure 1. Parallel ATA cable with 40-pin connectors for master and slave drives



The success of parallel ATA in the desktop and notebook markets is attributed to frequent performance enhancements and backward compatibility, with a constant goal of driving costs as low as possible. Since the introduction of parallel ATA, its data transfer rate has increased from 3 megabytes per second (MB/s) to 133 MB/s (Figure 2). ATA 100 and ATA133 have the headroom to handle the sustained transfer rate of today's7200 RPM HDDs because the interface only has to accommodate one drive at a time.

Figure 2. Data transfer rates for parallel ATA modes vs. the sustained transfer rate (STR) of HDD



The sustained transfer rate (STR) of a HDD is the rate at which the drive can transfer data sequentially from multiple tracks and cylinders to the disk interface. STR is dependent on multiple internal drive factors and it is different from the external transfer rate of the interface–ATA in this case.

This performance graph begs the question of why is there a need to change to a serial interface if ATA 100 can handle the requirements of today's desktop class (5400-RPM and 7200-RPM) HDDs. The answer concerns signaling voltage and data reliability. Parallel ATA data transfer is based on transistor-to-transistor logic (TTL) signaling. TTL signals define an 8-bit digital value, via a sequence of high and low voltage states, on pins 2 through 9 of the parallel port at a given point in time. TTL uses 5 V-tolerant, 3.3V signaling, which requires integrated circuits that can tolerate input signals up to 5 volts. It is becoming increasingly difficult to support the traditional 5-V TTL signal requirement because components are being fabricated with finer and more fragile lithographies.

With regard to data reliability, ATA uses cyclic redundancy checking (CRC) to verify the accuracy of the data signals transmitted between the host and HDD controller. However, ATA command signals are not checked with CRC, so they remain a potential error source.

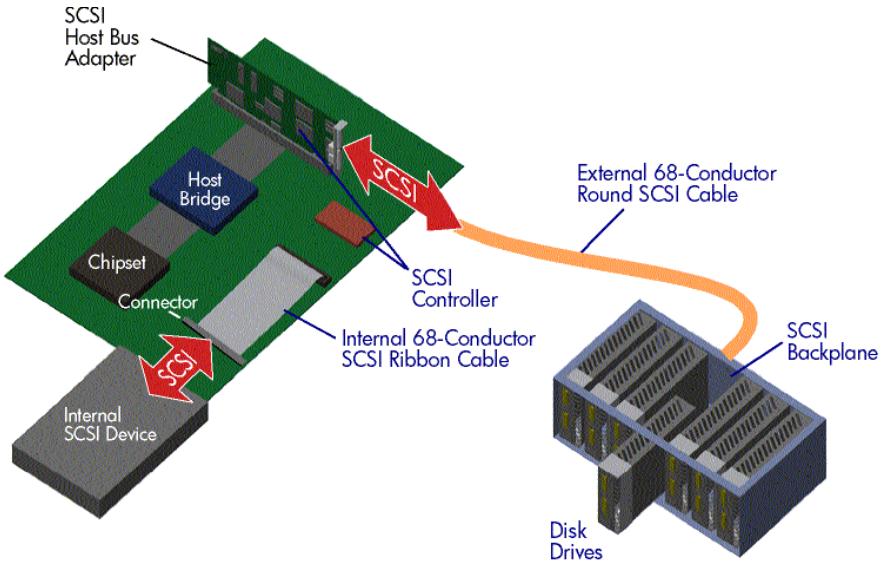
Going forward, it will be difficult to increase the speed of ATA beyond 133 MB/s due to the 5-V signaling requirement and the increased likelihood of command signal integrity issues.

Parallel SCSI

The SCSI specification was developed to provide a common interface that could be used across all peripheral platforms and system applications. The SCSI interface addresses a wider range of applications, such as Redundant Array of Independent Disks (RAID) storage, and has a broader command set than the parallel ATA interface. The success of SCSI as an I/O interface can be attributed to its performance, intelligence, and backward compatibility.

The SCSI system contains the SCSI controller (initiator), the SCSI bus (cable or backplane), and one or more target devices. The SCSI controller is the interface between the computer and the other devices on the bus. The SCSI controller may be built into the motherboard or housed on a SCSI host bus adapter (HBA) card in a PCI or PCI-X slot. Both configurations are shown in Figure 3.

Figure 3. SCSI components

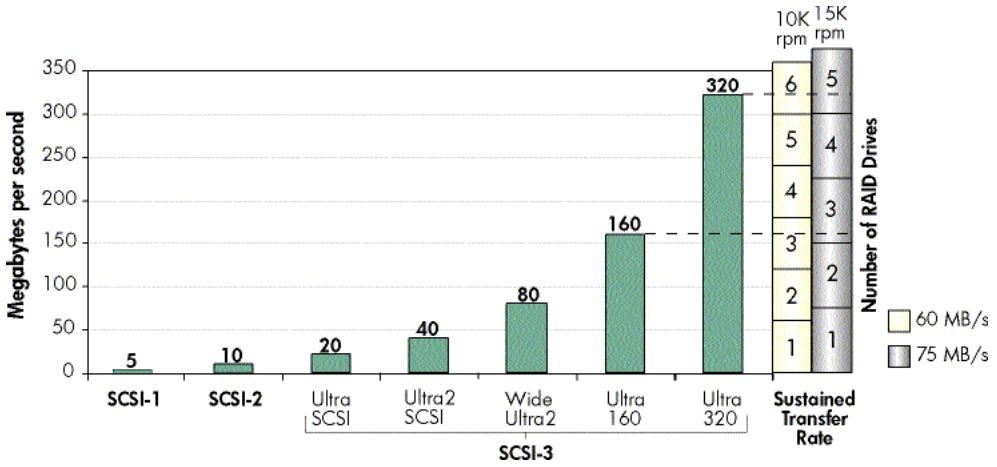


SCSI cables can connect up to 16 devices, including the SCSI controller. SCSI cables consist of 34 twisted pairs of multi-stranded flexible copper wires for a total of 68 conductors. SCSI devices inside the server are connected to the SCSI controller by using a 68-pin ribbon cable. The ribbon cable has a connector at each end and one or more connectors along its length. External SCSI devices are attached to the SCSI HBA by using a round 68-pin cable. Two sets of terminators, one at each end of the SCSI bus (not shown), prevent signal reflections within the cables.

Since 1981, there have been seven generations of the SCSI protocol. Each new generation has doubled the performance of the previous one (Figure 4). SCSI performance has ranged from an 8-bit, single-ended interface transferring data up to 4 MB/s (SCSI-1) to the latest 16-bit, low-voltage differential interface transferring data at 320 MB/s per channel (Ultra320 SCSI). The figure also shows the number of 10K RPM and 15K RPM drives that Ultra 320 SCSI can handle based on the

drives' sustained transfer rate (STR) during read operations. This is critically important for RAID implementations because RAID drives share the bandwidth of the interface.

Figure 4. Bandwidths of seven generations of SCSI compared to the sustained transfer rates of 10K and 15K RPM RAID drives



The development of Ultra320 SCSI was very challenging because it went beyond simply doubling the clock frequency of Ultra160 SCSI. In fact, the higher performance of Ultra320 SCSI was made possible by the implementation of several performance enhancements[1] , including:

- Read and write data streaming
- Quick arbitration and selection
- Flow control

Ultra320 SCSI also introduced two new technologies—pre-compensation and training—to minimize the negative impact of signal skew (slight signal delays from one wire trace to the next) and attenuation on signal integrity. Because the development of Ultra320 SCSI presented significant electrical engineering challenges, it is generally believed that Ultra640 SCSI cannot be deployed reliably without using new and expensive technologies. Therefore, the next performance increase will be 3 Gb/s with Serial Attached SCSI.

Serial drive technologies

To overcome the limitations of parallel I/O architectures, the server industry is migrating to serial I/O technologies. Serial I/O technologies have the potential to shrink form factors, lower power consumption, and extend I/O performance to meet the bandwidth requirements of a new wave of technology advances. Serial I/O drive technologies have the following features in common:

- Low-voltage differential (LVD) signaling
- Point-to-point connections
- 8b/10b encoding

LVD signaling

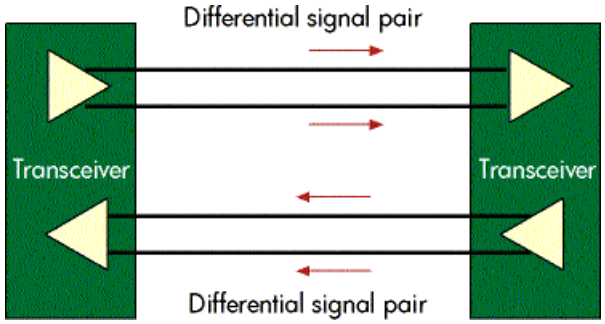
A major problem with parallel bus architectures is that as the signaling speed increases, signal skew and Inter-symbol Interference (ISI) can reduce the reliability and integrity of the bit signals. ISI is caused when a bit signal in a cable line is held at one voltage so long that the line becomes charged,

[1] For more information, read "Ultra 320 SCSI Maintains SCSI Prevalence" technology brief, document number TC021006TB.

much like a capacitor becomes charged. If a host device tries to send a short single-bit signal in the opposite direction by transitioning the voltage on the charged line, it is possible for the target device to completely miss the transition.

Serial technologies transmit signals in a single stream rather that in multiple parallel streams. Serial technologies incorporate a LVD signaling scheme that uses two pairs of data lines to transmit and receive low-voltage signals. The data is represented by the voltage potential between the two wires in each pair (Figure 5). Because it takes less time to apply low voltages to the wires, LVD signaling can occur at a much greater speed than in parallel architectures. The low voltage reduces the effects of capacitance, inductance, and noise. Noise sources tend to add the same amount of voltage to both wires, so the voltage difference between the wires remains the same.

Figure 5. LVD signaling



Point-to-point connections

The serial I/O architecture allows the design of switched point-to-point links between the host and multiple devices. This point-to-point architecture can support multiple simultaneous connections between the host and target devices. Each connection can be scaled by adding multiple links. As a result, point-to-point architectures provide significantly higher throughput than parallel architectures.

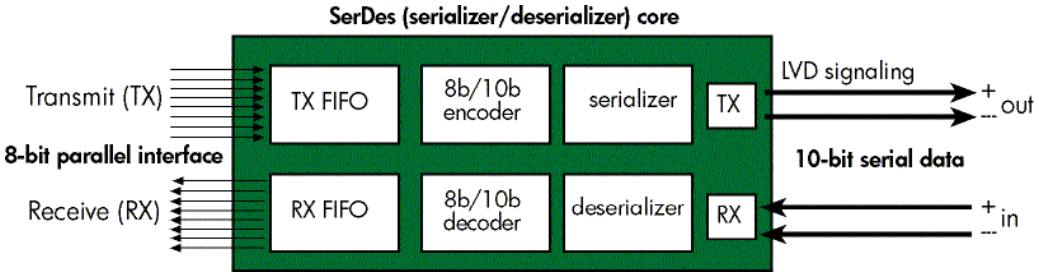
8b/10b encoding

With parallel bus architectures, the data and clock signals are transmitted along parallel wires from the initiator into the target device at a specific signaling rate. As the signaling rate increases, it becomes increasingly difficult to keep the data and clock signals aligned due to skew. In addition, signal integrity is degraded by the electrical noise that results from switching all data signals at the same time.

Serial architectures encode (embed) the clock signals into the data stream, thus eliminating the problem with aligning data and clock signals. Serial architectures require significantly fewer data lines to switch simultaneously, which reduces the introduction of electrical noise. As a result, serial signaling rates can be increased well beyond those attainable with a parallel bus.

Serial communication requires a device to convert parallel data into a serial bit stream and vice versa. This device, called a serializer/deserializer (SerDes), contains a parallel digital interface, First-In-First-Out (FIFO) caches, 8 bit/10 bit (8b/10b) encoder and decoder, a serializer, and a deserializer (see Figure 6). The 8b/10b encoder converts each 8-bit data byte to a 10-bit transmission character, which enables clocking information to be encoded into the data stream. Although this adds about 20 percent embedded overhead to the data stream, it eliminates the clock skew problem experienced by parallel architectures.

Figure 6. The SerDes core integrates 8b/10b coding and decoding logic



Serial ATA technology

SATA addresses the electrical signaling, cabling, and data robustness issues that inhibit enhancing the speed of parallel ATA beyond ATA 133. SATA offers lower signaling voltages, end-to-end data protection, hot-plug capability, and a reduced connector pin count (which results in a thinner cable that is easier to route).

The SATA specification is being released in three parts: SATA 1.0, SATA II Phase 1, and SATA II Phase 2. SATA 1.0 represents the transition from parallel to serial technology. SATA II Phase 1 adds features that address the needs of the lower cost server and network storage markets. SATA II Phase 2 increases the data transfer rate to 300 MB/s and adds other enhancements, which are described below.

SATA 1.0

The SATA 1.0 specification was released in August 2001 and accepted into ANSI T13. The specification focused on increasing the bandwidth and mitigating the design problems associated with the parallel ATA architecture. Serial ATA will be introduced with a maximum bandwidth of 1.5 Gb/s, or 150 MB/s factoring 20 percent encoding overhead (see "8b/10b encoding"). Other important changes are summarized in Table 1.

Table 1. Comparison of parallel ATA and SATA

	Parallel ATA	**Serial ATA**
Bandwidth	133 MB/s	150 MB/s
Volts	5V	250 mV
Number of pins	40	7
Cable length	18 in. (45.7 cm)	39 in. (100 cm)

SATA 1.0 is designed to replace parallel ATA, which means that it meets the requirements of desktops (non-hot plug). Serial ATA is not hardware compatible with legacy Ultra ATA; however, it is fully compliant with the ATA protocol and thus software compatible with existing ATA drivers. The key hardware components are described below.

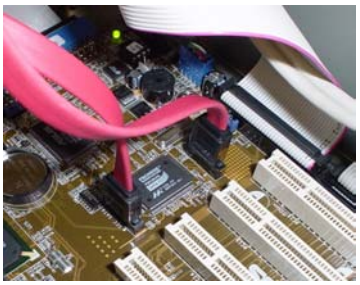
Key components

A SATA controller can be built into the motherboard (see Figure 7) or plugged into a PCI expansion slot. SATA controllers will have a number of ports, which can be deployed as either device connections or external connections.

SATA replaces parallel ATA's 40-pin connector and 18-inch long flat ribbon cable with a 7-pin connector (four signal lines and three ground lines) and a small diameter cable up to 1 meter long. The thinner cable improves both airflow and routing.

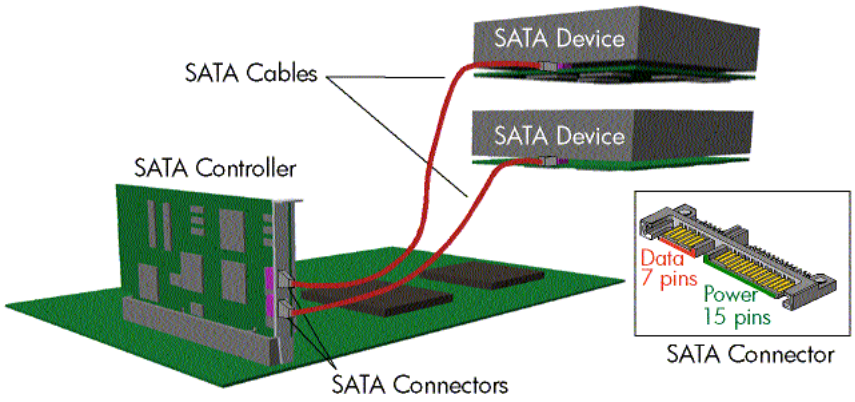
SATA 1.0 discards the parallel ATA Master/Slave concept and only allows one device per cable, which the system views as a master ATA device (Figure 8). These point-to-point connections allow each drive to communicate with the controller without having to wait for other data traffic to clear first.

Figure 7. SATA cables



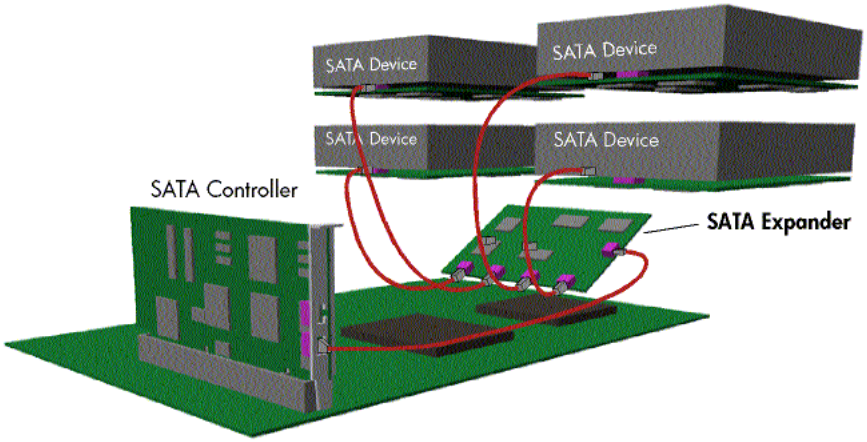
SATA also features a new 7/8-inch wide, 15-pin, single-row power connector. The power connector provides hot-plug capability, which allows a drive to be swapped out without powering down the entire machine.

Figure 8. Serial ATA Controller connecting with one device on each port to create point-to-point connections



The SATA controller can use an expander (low-cost switch) to "fan out" to multiple target devices from a single port (Figure 9).

Figure 9. SATA expanders allowing connectivity to a larger number of devices



Regardless of these useful features for PCs, the SATA 1.0 specification lacks support for some advanced features required by some server and network storage applications.

SATA II

SATA II is the first evolution of the SATA 1.0 specification. The SATA II specification is being developed in two phases. The SATA Phase I specification, which was released October 2002, improves the capabilities of SATA devices for server and networked storage applications. The SATA Phase II specification, which is expected to be released in the second half of 2003, will increase the signaling speed for all SATA market segments.

Performance enhancements

SATA Phase I features native command queuing, out-of-order execution and delivery, and data scatter/gather lists.

- Native command queuing enables a hard drive to take multiple requests for data from the processor and rearrange the order of those requests to maximize throughput. SATA II hard drives will be able to queue and execute requests without any assistance from the CPU.
- Out-of-order execution and delivery keeps execution resources as busy as possible. In the native command queuing model, this feature allows the last half of the data requested by a command to be delivered and executed before the first half of the data. If out of order data delivery within commands is desired, support for non-zero buffer offsets is required.
- A data scatter/gather list is a data structure that assists the direct memory access (DMA) engine in locating memory regions that comprise the complete transfer buffer. This assistance is beneficial because virtual memory mapping mechanisms may scatter the buffer across several noncontiguous, physical memory pages.

SATA Phase II increases the data transfer rate to 3 Gb/s, or 300 MB/s factoring 20 percent 8b/10b encoding overhead.

Backward compatibility

SATA II is not being developed in an open standards environment. It is controlled by the SATA steering committee (promoters). The advanced features in SATA II require updated operating system and driver support. SATA II devices will be backward compatible with SATA 1.0 but only at the SATA 1.0 data rate of 150 MB/s. SATA 1.0 devices will also work with Serial ATA II hosts.

SATA positioning and roadmap

Like its parallel predecessor, SATA will dominate the desktop market because it offers desktop reliability, functionality, and performance at a low cost. Due to the interface feature limitations and lower reliability of typical drive mechanisms, the primary use of SATA is in lower cost solutions such as PC desktops, notebooks, non-mission-critical server storage, and emerging markets such as consumer, gaming, and video recording.

SATA 1.0 HDD “bridged” solutions are expected to ship in late 2003. Bridged SATA HDDs have an additional application specific integrated circuit (ASIC) that converts or "bridges" SATA to PATA, which adds to the cost of the device and results in some performance limitations. The second-generation SATA HDDs that are targeted for end of 2003 will have integrated interface ASICs. SATA 1.0 controllers and HDDs are currently in development with product announcements slated for the first half of 2004. SATA optical devices will lag by possibly a year or more. SATA II Phase 1 features, such as hot plug and native command queuing, are targeted for second-generation devices beginning in the second half of 2003. SATA II Phase 2 devices with a signaling rate of 3.0 Gb/s are targeted for 2005.

Serial Attached SCSI technology

Serial Attached SCSI is a point-to-point disk interface that builds on the functionality of SCSI for the enterprise. The SAS architecture solves the clock skew and signal degradation problems experienced by parallel SCSI at higher signaling rates. SAS inherits its command set from parallel SCSI, frame formats from Fibre Channel, and physical characteristics from Serial ATA.

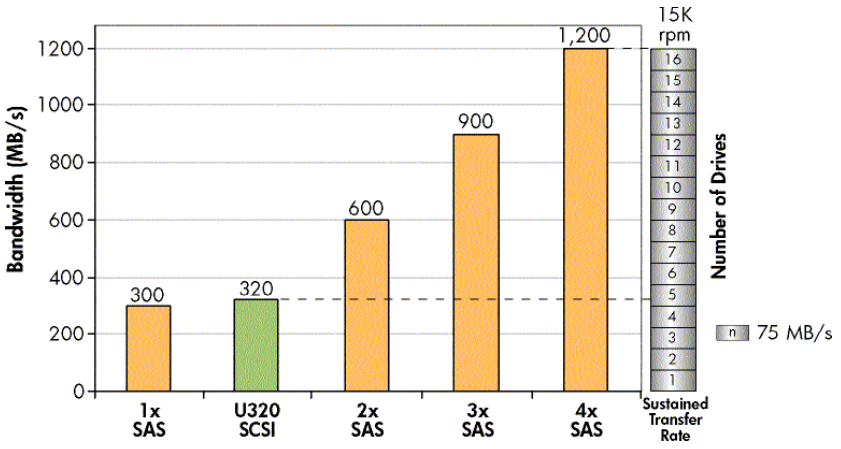
HP (Compaq), IBM, LSI Logic, Maxtor, and Seagate founded the Serial Attached SCSI Working Group in 2001. SAS is being developed in an open standards (ANSI T10) environment to ensure that the technology is an industry standard.

Performance

SAS achieves bandwidth as high as 300 MB/s per link and operates in full duplex mode, in which data, commands, or status information flow bi-directionally. The SAS interface allows for combining multiple links to create 1x, 2x, 3x, or 4x connections for scalable bandwidth. In contrast, Ultra320 SCSI has a half-duplex bandwidth of 320 MB/s per channel.

Figure 10 illustrates the bandwidths of SAS and Ultra320 SCSI compared to the bandwidth requirements of a number of 15K RPM drives. Notice that Ultra320 SCSI (per channel) can handle the STR of a maximum of four 15K RPM RAID drives while 4x SAS meets the requirements of up to 16 drives.

Figure 10. Bandwidths of SAS links and Ultra320 SCSI compared to the sustained transfer rate of 15K RPM HDDs



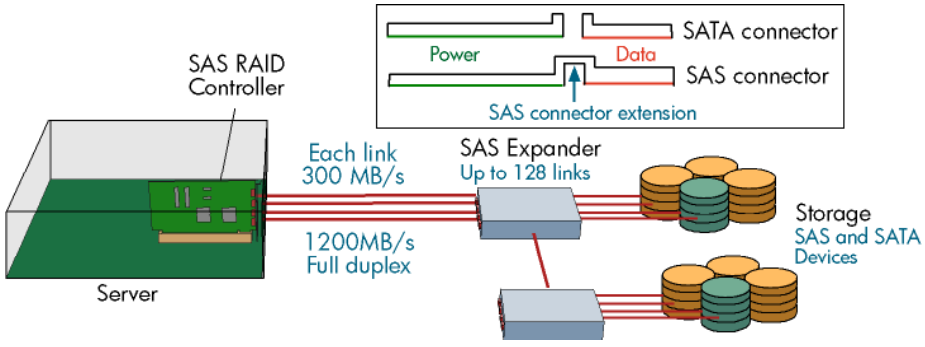
Key components

SAS controllers will be deployed as either internal device links or external links. All SAS devices can have one or more ports, with each port configured as a narrow (single link) or wide (multiple links) port. The SAS architecture supports up to 4,096 links, each link with a data transfer rate of 300 MB/s.

Expanders allow controllers to connect to a greater number of devices by providing up 128 physical links (Figure 11). These links may include SAS devices, SATA devices, other expanders, and other host connections.

The SAS interface allows devices to have two full duplex links, which are limited to no more that two simultaneous data transfers. Second generation SAS devices will have two ports to support redundant paths for high availability configurations.

Figure 11. Serial attached SCSI components



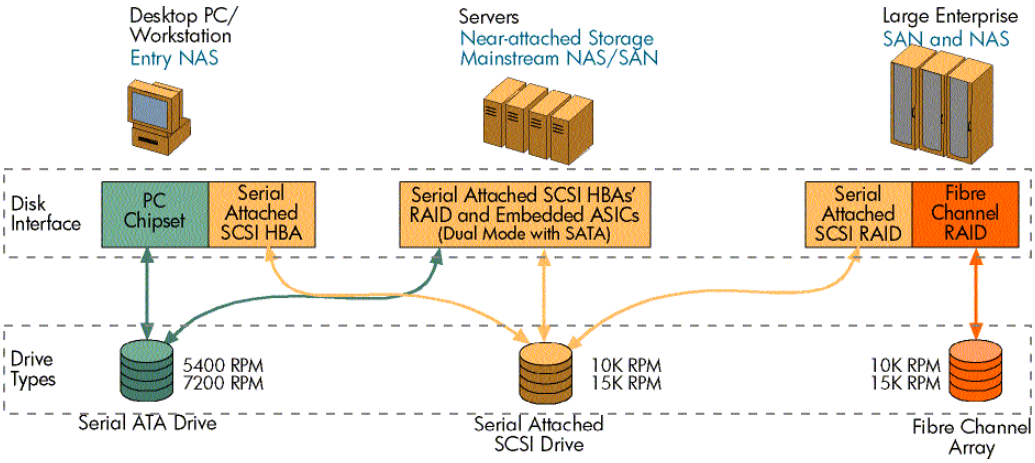
SAS/SATA compatibility

The SAS architecture enables system designs that deploy both SAS and SATA devices, a breakthrough for enterprise customers. This provides a broad range of storage solutions (Figure 12) that give IT managers the flexibility to choose storage devices based on reliability, performance, and cost. SAS and SATA devices will share the same physical device connector, with the exception of an

extension on the SAS connector. This extension will allow SAS to accept SATA device connections. However, SATA will not accept SAS device connections based on the current specification.

SAS supports three protocols to handle communications with various devices: Serial Management Protocol (SMP), Serial SCSI Protocol (SSP), and SATA Tunneling Protocol (STP). SMP is used to manage the point-to-point topology of expanders. SSP allows the controller to communicate with SAS devices and existing SCSI software. STP, on the other hand, allows SAS controllers to communicate with SATA devices. After power up or restart, the SAS domain goes through an initialization sequence to determine the proper protocol to communicate with the various types of devices.

Figure 12. The SAS architecture gives IT managers the flexibility to choose solutions



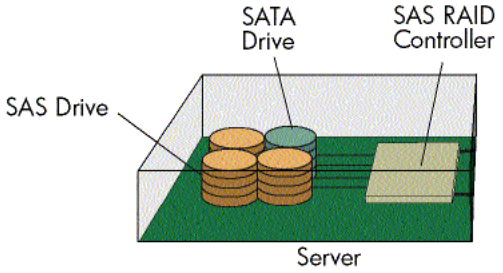
SAS topologies

Serial Attached SCSI enables highly scalable topologies—internal, external, or a combination of both—to give manufacturers and customers the flexibility to design and deploy a range of solutions.

Internal

Figure 13 shows a topology that can be used for internal RAID systems. The SAS RAID controller can support from two to eight internal HDDs, which can be either SAS or SATA devices. The small SAS connector will allow connections to 3.5-inch or 2.5-inch internal HDDs, enabling redundant storage configurations in dense server form factors like blade servers.

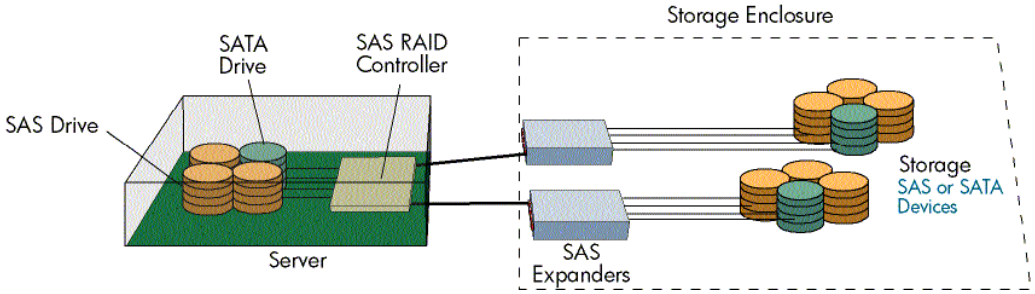
Figure 13. Topology for internal storage application



Internal/external with JBODs

The SAS interface will offer new levels of choice by enabling customers to plug either SAS or SATA drives into one backplane. This will give customers the flexibility to configure drive arrays with SAS, SATA, or both, enabling the use of enterprise devices and desktop devices in the same server or networked storage subsystem.

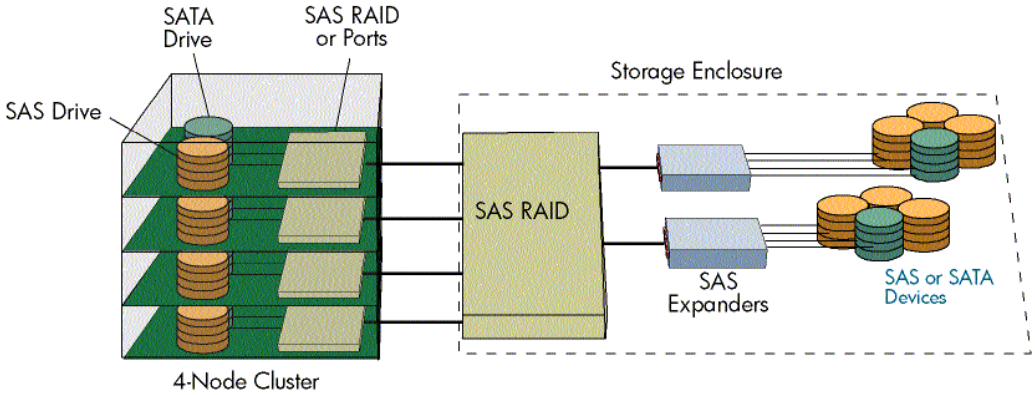
Figure 14. Topology for internal and external storage applications



Multi-node clusters

The highly scalable architecture enables topologies that support multi-node clustering for high availability (failover) or load balancing (Figure 15).

Figure 15. Topology for multi-node cluster applications



SAS positioning and roadmap

SAS compatibility with legacy parallel SCSI software will provide a migration path for existing SCSI users and allow SAS to replace SCSI as the dominant interface for DAS, NAS, and SAN systems.

SAS is not a fabric topology, so it is intended to complement, not replace, Fibre Channel in the SAN market. SAS compatibility with SATA will enable the deployment of multi-function systems that can provide high performance and high availability as well as low-cost-per megabyte solutions.

The SAS roadmap provides dual ported devices and increased data rates. SAS Emulators are targeted for the second quarter of 2003. First HDD and controller samples for development are targeted for the third quarter of 2003. SAS solution qualifications are targeted for the first half of 2004.

Conclusion

The benefits of SATA and SAS will extend from the desktop to the data center, including investment protection in SCSI software, higher HDD performance, longer cabling distances, smaller form factors, and greater device addressability. Because the SATA and SAS architectures use the same physical device connector, customers have the flexibility to design solutions that use both SAS and SATA devices. This flexibility is crucial for the adaptive enterprise.

HP is a leader in SATA and SAS technology development. HP is a contributor in the SATA technology development through the SATA Working Group. HP was also a founding member of the Serial Attached SCSI Working Group in November 2001.

Call to action

To help us better understand and meet your needs for ISS technology information, please send comments about this paper to: TechCom@HP.com.

TC030605TB, 06/2003